


Invited Review – A 5-year update on epigenome-wide association studies of DNA modifications in Alzheimer's disease: progress, practicalities and promise

A. R. Smith* , G. Wheildon*  and K. Lunnon* 

*University of Exeter Medical School, College of Medicine and Health, Exeter University, Exeter, UK

A. R. Smith, G. Wheildon and K. Lunnon (2020) *Neuropathology and Applied Neurobiology* 46, 641–653
Invited Review – A 5-year update on epigenome-wide association studies of DNA modifications in Alzheimer's disease: progress, practicalities and promise

In late 2014, the first epigenome-wide association studies of DNA modifications in Alzheimer's disease brain samples were published. Over the last 5 years, further studies have been reported in the field and have highlighted consistent and robust alterations in DNA modifications in AD cortex. However, there are some caveats associated with the majority of studies undertaken to date; for example, they are predominantly restricted to profiling a limited number of loci, are principally focused on DNA methylation, are performed on bulk tissue at the end stage of disease and are restricted to nominating

associations rather than demonstrating causal relationships. Consequently, the downstream interpretation of these studies is limited. Owing to recent advances in state-of-the-art cell profiling techniques, long-read genomic technologies and genetic engineering methodologies, identifying cell-type-specific causal epigenetic changes is becoming feasible. This review seeks to provide an overview of the last 5 years of epigenomic studies of DNA modifications in Alzheimer's disease brain samples and propose new avenues for future research.

Keywords: Alzheimer's disease, Epigenetics, EWAS

Alzheimer's disease: A global epidemic

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that accounts for around two-thirds of dementia incidence and is characterized by memory loss, difficulties with language and problem-solving and a reduction in cognition [2,3]. In 2015, dementia affected 46.8 million people globally, with this figure projected to exceed 74.7 million by 2030, and associated costs predicted to reach \$2 trillion per year [5]. Pathologically, AD is distinguished by the presence of extracellular

amyloid β (A β) plaques and intraneuronal neurofibrillary tangles (NFT) formed from hyperphosphorylated tau protein [7]. The characteristic spread of NFTs through the brain has been well documented and is described by Braak staging [9]. Other hallmarks of the disease include reduced synaptic capacity [10], neuronal cell loss [12] and the activation of glial cells [14]. Clinically, the pathological features of AD are reported to start years before the appearance of symptoms, with the preclinical lag from initial A β deposition suggested to be at least 15 years [17]. This delay in symptom onset, combined with a current lack of disease-modifying treatments and the predicted increase in disease incidence highlights the urgent need for further research identifying novel disease mechanisms.

*Correspondence: Katie Lunnon, University of Exeter Medical School, RILD Building Level 4, Royal Devon and Exeter Hospital, Barrack Rd, Exeter. EX2 5DW. UK. Tel: +44 (0) 1392 406758. E-mail: k.lunnon@exeter.ac.uk

The genomic era of AD

Defining the etiology of AD has proven more elusive than the characterization of the pathological and clinical features. Large-scale twin studies in AD have revealed a high rate of heritability (58%–79% [18]), and thus, over the last two decades, genomic studies have been at the forefront of the AD research field. Although it is well established that autosomal dominant mutations in genes that are part of the A β production pathway, such as amyloid precursor protein (*APP*) and the presenilin genes (*PSEN1*, *PSEN2*), result in familial AD (FAD) [19], which is typically early onset (EOAD), this only accounts for 1–6% of diagnoses [21]. In the case of sporadic AD, determining the genetic contribution to AD aetiology was first explored through genome-wide association studies (GWAS), which in recent years have been combined into meta-analyses. To date, the largest two such studies were published only last year [22,23]; Kunkle and colleagues nominated 25 risk variants in ~95,000 individuals, whilst Jansen *et al.* identified 29 risk loci in a study exceeding 450,000 individuals, as they included an AD-by-proxy phenotype. Across all studies, the gene that confers the greatest risk for disease still remains the *APOE* gene, which was first reported over 20 years ago; a single copy of the $\epsilon 4$ allele confers a two- to threefold higher risk, which rises to 12-fold in individuals with two alleles [25]. In recent years, it has been shown that AD has a considerable polygenic component [28], whereby many common variants of small effect size contribute to the heritable risk. The summative effect of all the known, and indeed yet to be identified, common variants is approximately 33%, including *APOE* $\epsilon 4$ [29]. This has led to the development of polygenic risk scores (PRS) for AD, that being a composite score of the genetic loading for disease. Although reports vary, Escott-Price *et al.* [28] reported a prediction accuracy of 78.2% when the PRS, *APOE* genotype, age and sex were used as predictors. Exome or whole genome sequencing projects have also nominated a handful of loci with modest effect sizes; however, these are relatively rare variants [30]. Interestingly, recent work has highlighted an increasing number of somatic single nucleotide variations (SNVs) in AD brain with age. These mutations were enriched in genes in pathways known to contribute to the hyperphosphorylation

of tau [31]. Therefore, although genomic studies in AD have given considerable insight into disease etiology, a large proportion of disease risk remains unaccounted, suggesting that other molecular processes may contribute to the burden of disease.

Moving to the next dimension: genome-scale studies of DNA modifications in AD

In recent years, focus has shifted onto the potential contribution of epigenetic mechanisms to AD etiology. Epigenetic mechanisms principally act through DNA and histone protein chemical modifications, thereby altering chromatin structure, to regulate gene expression in a reversible manner without altering the DNA sequence (Figure 1) [32]. In addition, noncoding miRNAs offer another level of epigenetic control by post-transcriptionally repressing gene expression (33). These processes are discussed in further detail in Box 1. DNA modifications, namely DNA methylation, are the most studied epigenetic modification in the context of AD and are the focus of this review.

Exploring the role of epigenetic processes in AD is by no means recent, with initial empirical studies dating back over 20 years. However, these were largely limited to profiling candidate genes, or assessing global DNA modifications in small numbers of samples (as reviewed in [34]). It was only owing to more recent advances in genomic technology that 5 years ago, the first epigenome-wide association studies (EWAS) of DNA modifications were published, which used the Illumina Infinium Methylation 450 K BeadChip Array (450 K array) [35,36], with more studies appearing on a year-by-year basis (Figure 2). Those first large-scale EWAS were illuminating as although different cohorts, tissues and sample sizes were used in the two studies, a number of differentially methylated positions (DMPs) were common to both studies including *ANK1*, *CDH23*, *RHBDF2* and *RPL13* [37]. The Lunnon *et al.* study represented a cross-tissue analysis of DNA modifications associated with Braak stage, as a measure of disease severity, in ~120 individuals utilizing three cortical brain regions (entorhinal cortex, superior temporal gyrus and prefrontal cortex) in addition to cerebellum and premortem blood [35]. They particularly focussed on the *ANK1* gene, demonstrating consistent hypermethylation across the cortex, with no association in the cerebellum or blood. Interestingly, the two *ANK1* DMPs

identified were shown to be consistently differentially methylated in AD cortex in three independent data sets, including one that used an alternative technology (pyrosequencing) to examine an extended region of the gene. The study by De Jager and colleagues only profiled the prefrontal cortex but leveraged on the power from >700 individuals to identify 71 plaque-associated DMPs, of which 12 were validated in the Lunnon data set [36]. The study highlighted in particular seven differentially methylated loci (*ANK1*, *CDH23*, *DIP2A*, *RHBDF2*, *RPL13*, *SERPINF1*, *SERPINF2*) as they also showed altered gene expression. One of the 12

validated DMPs in this study was shown to reside in the *HOXA* gene cluster. More recently, another EWAS, which leveraged on the validation data from the Lunnon *et al.* data set, highlighted a 48 kb region within this gene, consisting of 208 probes that showed hypermethylation in AD prefrontal cortex [38]. The study initially identified 10 genome-wide significant DMPs in the PFC in ~150 donors that were robustly associated with Braak stage, including one probe in the *HOXA3* gene. A subsequent analysis to identify differentially methylated regions (DMRs) consisting of multiple neighbouring probes identified six closely located DMRs

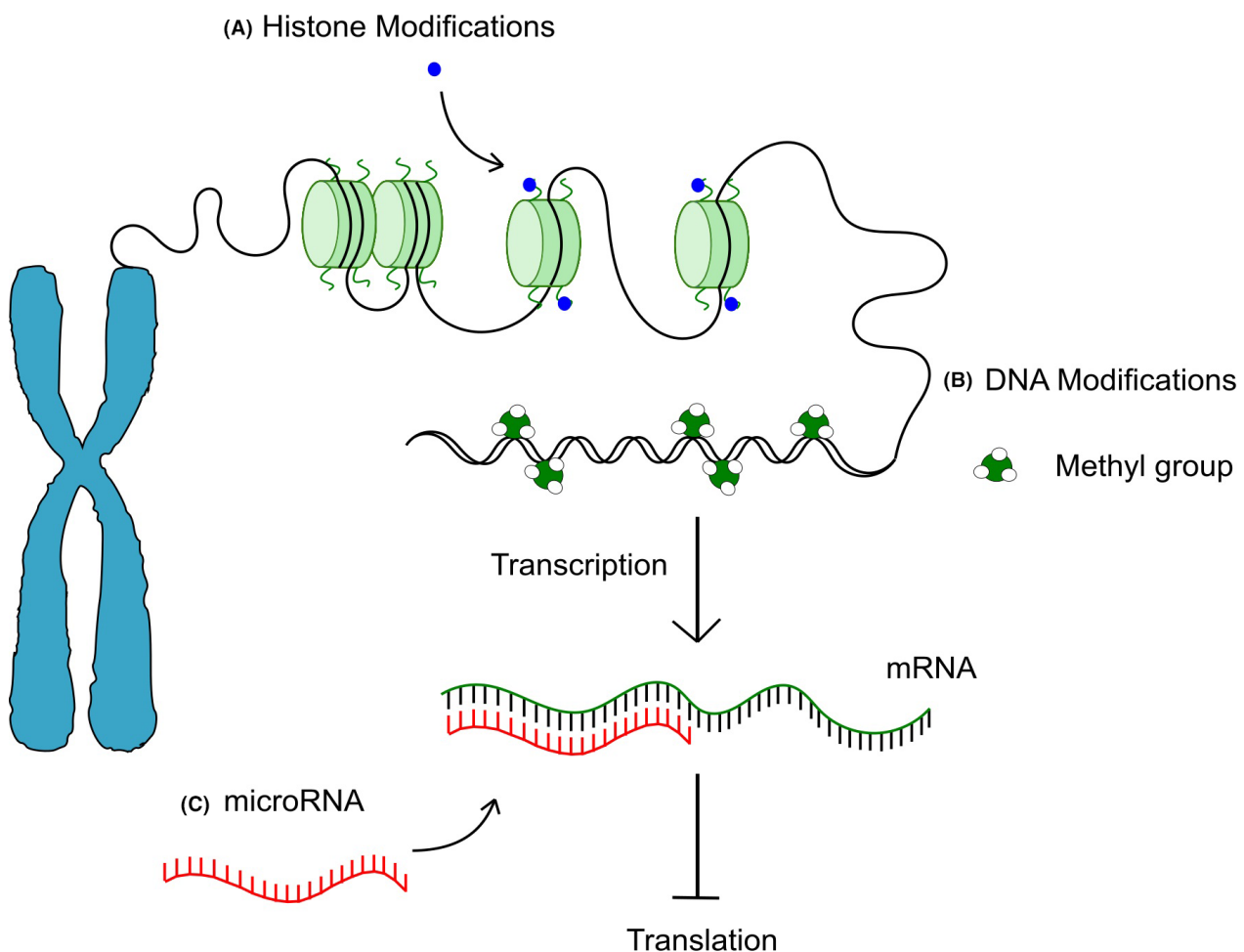


Figure 1. An overview of different epigenetic mechanisms regulating gene expression. (A) The regulation of chromatin structure through post-translational modifications to histone proteins, including acetylation, methylation, sumoylation, ubiquitylation, citrullination and ADP-ribosylation. (B) The addition of chemical tags to cytosine usually in the context of a CpG dinucleotide, resulting commonly in DNA methylation (5mC), as well as its derivatives 5hmC, 5fC and 5caC. (C) The regulation of transcription via small RNA molecules, such as microRNAs.

Box 1**Overview of epigenetic mechanisms****DNA modifications**

DNA methylation is the most stable and best characterized epigenetic mechanism, which occurs when DNA methyltransferases (DNMTs) transfer a methyl group to a cytosine residue, usually in a CpG dinucleotide, forming 5-methylcytosine (5mC) [1]. Traditionally, this process was thought to result in transcriptional silencing, by blocking the binding of transcription factors [4] or recruiting methyl-binding proteins, which in turn recruit transcriptional repressing complexes that remodel chromatin structure [6]. However, recent evidence suggests that the relationship between DNA methylation and gene expression is more complex and context dependent; gene body methylation has been reported to correlate with increased gene expression (8), whilst intragenic methylation has been shown to modulate alternative splicing [11]. Although 5mC is the most studied DNA modification, other cytosine modifications have been reported in recent years; 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) were originally thought to be intermediates in the demethylation pathway. However, 5hmC in particular has been increasingly studied as an independent epigenetic mark in the context of neurobiological phenotypes owing to its higher levels in the brain [13] and enrichment in neurons [15] and genes encoding synaptic proteins [16].

Histone modifications

Chromatin compaction and therefore transcriptional ability is regulated by conformational changes to histone proteins, including modifications such as methylation, phosphorylation, acetylation, ubiquitination, sumoylation and adenosine diphosphate-ribosylation as well as other post-translational modifications to the amino acids that make up histone protein structure [20]. These histone 'tail' modifications regulate the level of condensed state chromatin (heterochromatin) and relaxed chromatin (euchromatin) denying or allowing access to transcription start sites (TSS) and promotor regions within the DNA to transcription factors and other necessary cellular machinery required for transcription. Acetylation and methylation are the most widely studied histone modifications. Histone acetylation appears to be a ubiquitous marker of gene expression (24). The second most studied histone modification is methylation, which has differing effects on gene transcription based on which amino acid is methylated and the number of methylation groups per amino acid.

MicroRNAs

Micro-RNAs (miRNAs) play a critical role in gene regulation. Currently, 2,654 mature human miRNAs are listed in the latest release [22] of miRBase [26]. miRNAs are small (19–26 nucleotide) RNAs that regulate gene expression post-transcriptionally by binding to their target mRNAs and inhibiting translation or by cleaving them [27]. There is considerable promiscuity in miRNAs, with multiple predicted targets for a given miRNA.

within the *HOXA* gene region, which upon further inspection showed an extended region of disease-associated hypermethylation spanning >48 kb [38]. Further EWAS using the same technology and principles have provided additional evidence for DNA methylomic alterations in AD; Watson and colleagues identified 479 diagnosis-associated DMRs by profiling the superior temporal gyrus in 34 control and 34 AD samples. These regions were enriched for genes involved in neuronal function and development, and a subset of the DMRs were independently associated with aging [39].

A study by Mano *et al.* [40] attempted to address the issue of tissue heterogeneity by profiling the neuron-specific DNA methylome of post-mortem brain samples (inferior temporal gyrus) from 30 AD patients and 30 age-matched normal controls. They identified eight DMRs consisting of 36 statistically significant probes. They highlighted that protein levels of BRCA1 (their most significant DMR) were increased in the cytoplasm of neuronal cells in AD brains. A more recent study by Gasparoni *et al.* [41] also looked to address the issue of tissue heterogeneity, this time by profiling both bulk

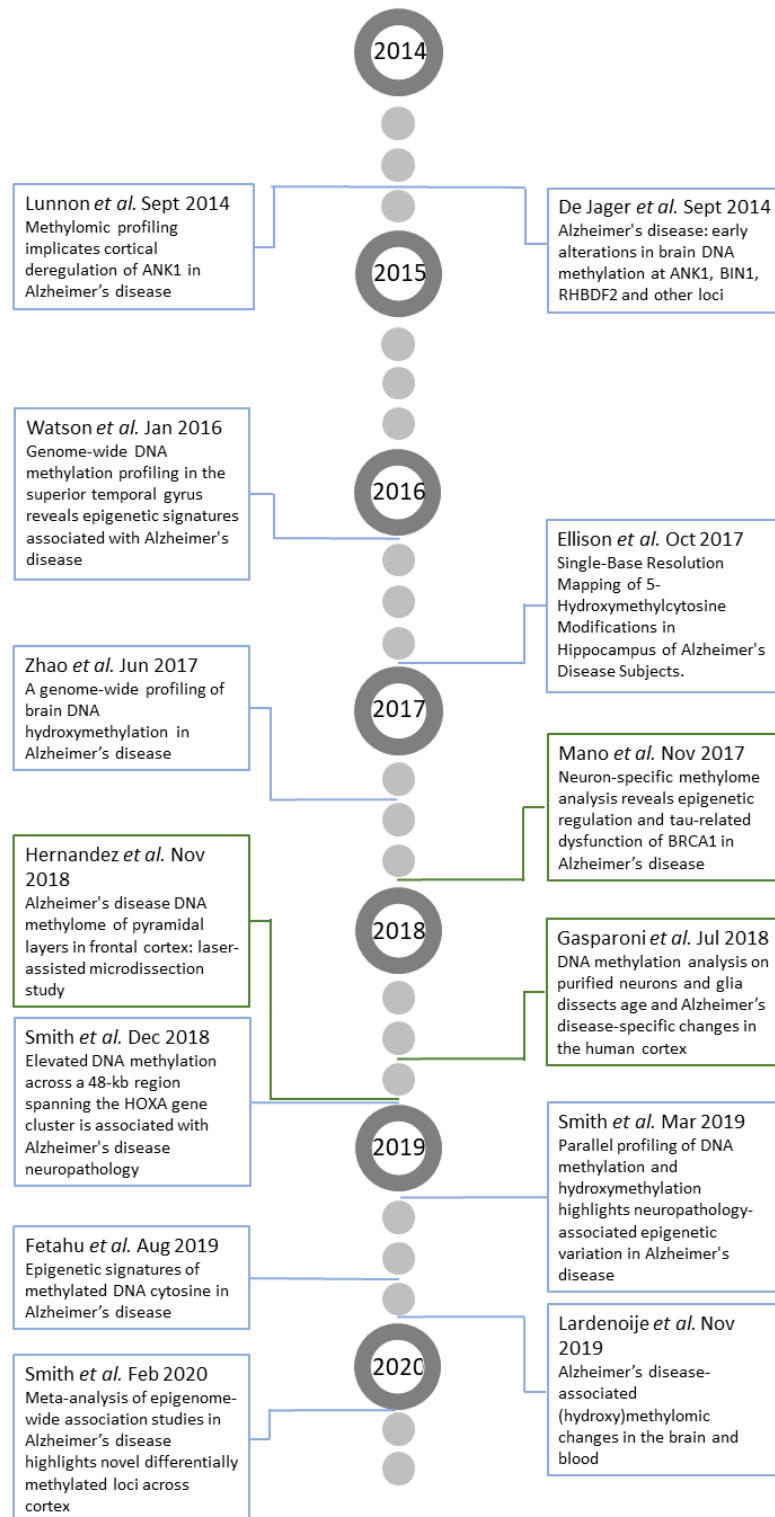


Figure 2. A timeline depicting the progress of epigenome-wide association studies of DNA modifications in AD brain that have taken place over the last 5 years. Studies of bulk tissue are shown in blue, whilst those studies that have focused on cell populations are shown in green.

tissue (matched temporal cortex and frontal cortex from 65 donors with varying degrees of AD pathology) and sorted neuronal and non-neuronal (glia) nuclei from the occipital cortex in 31 independent donors. No loci reached genome-wide significance in the bulk tissue, and furthermore, their top ranked loci showed no overlap with the study by Lunnon *et al.* [35]. Interestingly however, by performing methylomic profiling independently in neuronal and glia cells, they were able to identify significant Braak-associated DMPs, demonstrating that disease-associated methylation changes in *HOXA3* and *ANK1* correspond to specific changes in neurons and glia respectively [41].

One caveat of all of the aforementioned EWAS of DNA methylation in AD is that they have been performed on bisulfite-treated DNA, which is unable to distinguish between two different DNA modifications (5-methylcytosine [5mC] and 5-hydroxymethylcytosine [5hmC]) (see Box 1) and, as such, the studies are actually reporting a sum of both modifications. A recent adaptation to the way the DNA is processed means it is possible to separately quantify 5mC and 5hmC. Smith and colleagues used this method in conjunction with the 450 K array in 96 entorhinal cortex samples to nominate a number of differentially hydroxymethylated positions (DHPs). One particular highlight of the study was the finding that previous reports of disease-associated *ANK1* DNA hypermethylation were, in fact, underestimates as it was confounded by DNA hypohydroxymethylation [42]. More recently, Lardenoije *et al.* also used the same approach to quantify DNA methylation and hydroxymethylation in the middle temporal gyrus of 45 AD and 35 control donors. In addition to exploring alterations in 5hmC, they highlighted a DMR in the *OXT* gene consisting of 10 neighbouring CpG sites, which was hypomethylated in AD. Interestingly, the authors also reported significant DNA methylation changes in blood associated with conversion to AD dementia in this same genomic region in a separate preclinical cohort of 42 individuals [43]. Two other studies that have been published to date have also explored 5hmC levels in AD brain samples but using alternative technologies. Zhao and colleagues [44] used a selective chemical labelling technique to enrich for 5hmC, and then sequenced the captured libraries from a cohort of 30 individuals with either no dementia, mild cognitive impairment (MCI) or AD. They nominated 517 plaque-associated and 60 NFT-associated differentially hydroxymethylated regions (DhMRs) and showed that

5hmC is enriched in genes with neurobiological functions. Although promising, the study did feature low sequencing resolution, so it was unable to completely distinguish between 5mC and 5hmC [44]. Another study published that year utilized reduced representation hydroxymethylation profiling (RRHP) to analyse 5hmC levels in the hippocampus of three AD cases and two age-matched controls. This showed an enrichment in a number of relevant cellular pathways [45]. The key advantage of this study is that this technology allowed profiling of 5hmC at >2 million genomic locations, which represents >4 times more coverage than the 450K array; however, given that AD is a very heterogeneous disease, the very small sample size limits the utility of the data.

One striking observation of all AD DNA methylation EWAS performed to date is the considerable overlap in loci nominated across studies. This was particularly evident in a recent AD EWAS meta-analysis, which combined data from six independent AD cortical EWAS totalling >1,400 unique donors, many with multiple brain regions available [46]. The study identified a number of loci that showed significant differential methylation across the cortex. These loci showed a similar pattern of methylation change across all cohorts and cortical tissues and were validated in additional independent cohorts (Figure 3).

Profiling of heterogeneous tissue: Problems, practicalities and possibilities

Despite the relative infancy of epigenomic studies in AD, the field is rapidly building momentum and there is now considerable evidence demonstrating robust and reproducible genome-wide changes in various epigenetic modifications in AD. However, there are a number of caveats of the EWAS in AD undertaken to date; for example, they are largely restricted to profiling bulk tissue at the end stage of disease, at a specific set of methylation sites and cannot infer causal relationships. However, owing to recent advances in technology, it is becoming feasible that these could be addressed in the next wave of studies.

The issue of mixed pathologies and inconsistent diagnostic criteria

One particularly pertinent issue when studying post-mortem human brain tissue is that many individuals

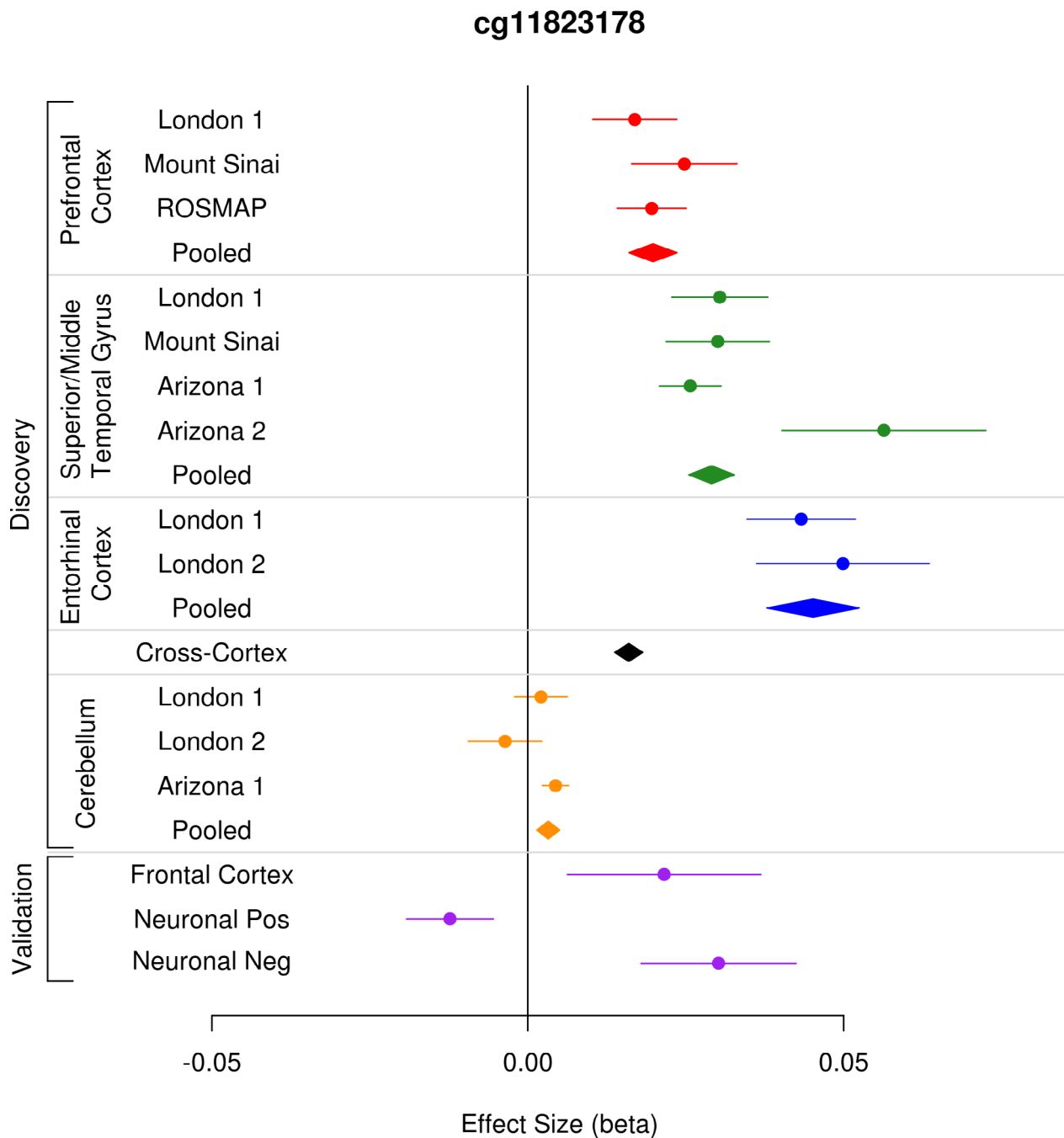


Figure 3. A forest plot depicting DNA modification differences between individuals with high AD pathology and no AD pathology for a site within *ANK1*. *ANK1* has consistently been reported to show hypermethylation in AD cortex across many independent studies of DNA modifications. A recent meta-analysis of all available EWAS of DNA modifications included cortical data from >1,400 unique donors from across six different cohorts for discovery, many with multiple brain regions available. The forest plot highlights that for probe cg11823178 there is consistent hypermethylation in the six discovery cohorts across the cortex (black), specifically in the prefrontal cortex (red), temporal gyrus (green) and entorhinal cortex (blue), which was not seen in the cerebellum (orange). In an independent cohort of validation samples (purple), the pattern of pathology-associated hypermethylation was seen in the frontal cortex and seemed to be driven primarily by non-neuronal cells, rather than neuronal cells. Cohorts and tissues are shown on the Y-axis, whilst effect size (beta), representing the difference in DNA methylation between Braak stage 0 (no NFT pathology) and Braak stage VI (most severe NFT pathology), is shown on the X-axis. Taken from Ref. [46]

have mixed pathology. As such, stratifying 'AD' cases from 'controls' based purely on clinical or neuropathological criteria for AD could result in the inclusion of individuals with other co-existing neurodegenerative conditions in both groups. Indeed, a recent study profiling *ANK1* DNA methylation in the cortex in various different neurodegenerative diseases highlighted hypermethylation of this gene in individuals with vascular dementia (VaD) and dementia with Lewy bodies (DLB), only when individuals had co-existing AD pathology [47]. One mechanism to control for this would be to only include individuals in a study with pure 'AD' pathology and no other neurodegenerative conditions. However, given that there are a limited number of such brain samples available, this would considerably limit the power of the study. Another option is to analyse individuals with mixed pathologies, and then control for this during the data analysis.

Techniques to address sample heterogeneity in AD methylomic studies

Epigenetic modifications are, by their nature, cell-type specific, allowing the development and programming of specific cells that fulfil key roles in the body. In addition, these modifications are known to be dynamic, able to change a cell's transcriptional profile to adapt to the cellular environment. It is therefore likely that distinct disease-associated epigenetic changes will occur in different cell types. Furthermore, it is known that there are alterations in the proportion of different cell types in AD brain, with a loss of neurons and gliosis [48]; given that most studies to date have used bulk tissue, this therefore poses a considerable limitation. In an attempt to circumvent this, most of these studies have attempted to bioinformatically correct for cellular proportions in EWAS data using methods that can estimate cell proportions, which can either be based on well-annotated reference data sets in separated cell populations (reference-based) or can be reference-free. Although these methods provide reassurance that loci nominated in the EWAS are not directly attributed to a shift in cell proportions, they are still not able to discriminate which cell type is driving a specific epigenetic signal, which greatly limits the interpretation of findings. Aside from this, if epigenetic changes in one cell type oppose epigenetic changes in another cell type, then no total difference in epigenetic levels will be seen in bulk tissue, meaning that many

important disease-associated changes may be missed with this approach. To date, one EWAS in AD has attempted to identify cell-specific changes; as previously described, Gasparoni and colleagues [41] profiled DNA methylation using the 450 K array in fluorescence-activated cell sorted (FACS) neuronal (NeuN+ nuclei) and non-neuronal (NeuN- nuclei) populations in post-mortem occipital cortex from 31 individuals and, importantly, were able to show the cellular origin (neuronal/non-neuronal) of many previously reported EWAS loci, including *HOXA3* (neuron) and *ANK1* (glia). Although this study has been illuminating, given the relatively small sample numbers and the use of the occipital cortex, a brain region affected late in disease, these results require further replication. In addition, there are also limitations to this approach as it based on the expression of a single antigen (NeuN) and so it is not possible to determine whether the non-neuronal findings originated from microglia, astrocytes, oligodendrocytes or other cell types. Looking to the future, a wider range of antibodies is required, which can detect nuclear proteins only expressed in a specific cell subtype and that work efficiently in frozen post-mortem tissue.

Aside from FACS, laser capture microdissection (LCM) is another methodology that offers potential for identifying cell-type-specific epigenetic changes in AD. This technique also has some key advantages over FACS as it allows the isolation of particular cell types within certain structures of the brain. Furthermore, in the context of AD, LCM would allow the isolation and comparison between NFT-bearing neurons and unburdened neurons within the same sample, which could allow the identification of epigenetic changes prior to the presence of pathology within a cell. LCM has successfully been used in one AD EWAS to date; Hernandez *et al.* [49] isolated pyramidal layer cells in 32 frontal cortex post-mortem brain samples, including 18 AD donors, before profiling on the 450 K array and highlighted differential methylation in several previously nominated loci, including *HOXA3* and *BIN1*, in addition to demonstrating an enrichment of genes associated with oxidative stress and synapsis. However, this study was limited to a small sample number and just explored this single population of cells. As LCM is more labour intensive than FACS and yields of DNA are usually lower, this means it is currently not suitable for large cohort studies exploring different cell types.

Improvements in methodologies to allow unbiased DNA modification profiling

The majority of EWAS of DNA methylation to date have been undertaken using the 450 K array. However, one issue with this approach is the limited number of methylation sites that are profiled (485, 577). Although its successor, the EPIC array, provides nearly twice the coverage (>850,000 sites), this array like its predecessor is mainly focussed on cytosine methylation predominantly in CpG dinucleotides. Other methodologies with wider coverage that are available and could be applied for AD EWAS include whole genome bisulfite sequencing (WGBS), where all DNA is sequenced, and reduced representation bisulfite sequencing (RRBS), where CG-rich regions of the genome are enriched and sequenced. These techniques have largely not been used in human AD EWAS due to the availability of the 450 K/EPIC arrays; however, they are being applied in animal models that do not have the luxury of a cost-effective microarray platform being available for profiling DNA methylation at present. Recent adaptations to these sequencing methodologies mean that single-cell (sc) WGBS and scRRBS protocols, amongst others, have been developed and are being applied in other fields (for a review of technologies, please see [50]). The ability to profile epigenetic variation at the level of the single cell holds particular promise for a breakthrough in our understanding of epigenetic processes in AD.

One issue with bisulfite-based methods is that this has been shown to introduce sequencing bias [51]. Recent developments in third-generation (long-read) sequencing technology are also providing new avenues for epigenetic research. Aside from their ability to sequence considerably longer read lengths than next-generation (short-read) methodologies [52], there are a number of other key advantages to these technologies; they are able to profile native DNA, meaning there is no potential PCR bias and that it is possible to simultaneously perform genetic and epigenetic analysis, including a range of DNA modifications, including those to other bases [53,54]. As previously described, a number of recent EWAS have used bisulfite treatment and oxidative bisulfite treatment of DNA in parallel prior to analysis via sequencing or on the 450 K array to quantify 5mC and 5hmC independently. However, these methods do not quantify other cytosine modifications, namely 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (see

Box 1), where their functional relevance is controversial as they are known to be very lowly expressed in brain tissue [55]. Aside from long-read third-generation sequencing, methylase-assisted bisulfite sequencing (MAB-seq) is another method that allows base resolution mapping of 5fC and 5caC, by enzymatically converting unmodified cytosine to 5mC [56]. A recent study used this technique alongside oxidative bisulfite sequencing in induced pluripotent stem cells (iPSCs) models of AD to allow the simultaneous quantification of four cytosine states (5mC, 5hmC, 5fC/5caC and unmodified cytosine) [57]. They identified different DNA modification signatures of neuronal differentiation and also explored these in three post-mortem brain samples (1 control, 2 AD) as well as validating the 5mC profile from a previously published EWAS of AD [36]. There is also the possibility that this technology can be applied to study 5fC and 5caC at the level of the single cell [58].

Establishing epigenetic causality and time course of changes by applying novel methodologies

Possibly the most limiting factor in interpreting epigenetic associations with disease outcomes is the question of whether nominated loci are causal or are secondary to the disease. In recent years, computational approaches, such as two-step Mendelian randomization, have been used in the context of other phenotypes to address this issue [59]. In this approach, a genetic proxy for DNA methylation is used to interrogate the causal relationship between DNA methylation and outcome. However, the approach itself has some limitations, for example, it is only feasible if a genetic proxy can be identified and large sample numbers are required [60].

Experimentally, genetic engineering techniques could be utilized to investigate epigenetic causality. The clustered regularly interspersed short palindromic repeats (CRISPR) system uses RNA-guided Cas9 nucleases and is mainly used to introduce DNA breaks which can be repaired through homologous recombination, insertion-deletion (indel) mutations or with a vector carrying a desired mutation [61]. Recent modifications of this technology have allowed the alteration of DNA methylation at a selected locus, by using a Cas9 protein that is fused to the enzymatic domains of the DNA methyltransferase (DNMT) or ten-eleven translocation (TET) enzymes, which can then add or remove methyl groups at a specific locus, respectively [62]. This would

allow target-specific epigenetic editing, and by measuring the downstream molecular and cellular consequences, could in turn shed light on whether EWAS-nominated loci are causal in disease. Although these techniques could be undertaken in a commercial human brain cell line, such as the SHSY-5Y (neuronal) and SV40 (glial) lines, its application in induced pluripotent stem cells (iPSCs) from patients particularly merits investigation in the context of AD. This would allow the study of whether iPSCs harbouring specific genetic backgrounds (i.e. a high polygenic load for AD) have different changes in phenotype as a result of the epigenetic editing, compared to samples with a low genetic load. In addition, as iPSCs can be transformed into any cell type, it would also be one to investigate whether altering DNA methylation in different cell types elicits distinct phenotypes, which is an important point when considering translational potential. One recent advance in iPSC technology is the generation of brain organoid models, which represent 3D cultures of multiple CNS-relevant cell types generated from iPSCs. The use of epigenetic editing in these systems has the distinct advantage in that multiple brain cell types are present, and they better resemble the environment in the brain. However, there are still caveats to this approach, for example, their small size limits their applicability for modelling an adult brain and they have slow growth, adding considerable cost [63].

Epigenetic changes are specific to a particular cell type, and thus, EWAS to identify disease mechanisms need to be undertaken in an appropriate tissue (e.g. the cortex), as studies in non-affected tissues (e.g. the cerebellum) deliver little mechanistic insight [35]. Although some studies have explored DNA methylation profiles in AD blood samples, the primary focus of those studies has been on identifying biomarkers for disease, rather than elucidating disease processes. One particular limitation of human epigenetic studies is that these are thus limited to analyses of post-mortem tissue at a single (end-point) stage of disease, and so it is not possible to explore the temporal pattern of methylation changes with respect to the onset of pathology. Numerous rodent models of AD are available; however, these have traditionally been murine models of familial AD or frontotemporal dementia (FTD) expressing humanized genes bearing disease-causing mutations. Although these allow the mapping of temporal changes in DNA methylation with respect to pathology, they do not

allow the identification of causal epigenetic mechanisms. However, a recent study has shown pathology development and spread in a mouse model expressing the human *MAPT* (tau) gene in the absence of any disease-causing mutations after it was inoculated intracerebrally with tau extracts from a post-mortem AD brain sample [64]. This represents an important step forward in the field, as profiling epigenetic variation in this (sporadic) model could help to identify causal epigenetic processes in sporadic AD.

Concluding points

Overall the study of epigenomic mechanisms in AD has advanced considerably over the last 5 years. In an emerging world of advanced laboratory techniques, AD is becoming a data-rich area of research at other molecular levels. It is particularly important that DNA methylomic studies in AD are integrated with other 'omic' data sets, for example, those quantifying genetic variation, histone modifications, miRNAs, transcription and translation. When this can be achieved at the level of the single cell, this will represent a particular advancement in the field. One challenge in this respect is in our ability to analyse and integrate large genome-wide data sets and advances in computational methods will be the turning point in facilitating the interpretation of epigenomic data in the context of AD.

One of the most exciting aspects of identifying disease-associated epigenetic changes is that they are potentially reversible and so could represent viable therapeutic targets. However, even in the genomics field, despite numerous robust findings over the last 10–15 years, there has been very limited translation to novel pharmacological compounds for AD. In reality, there are currently no disease-modifying treatments for AD and many clinical trials have been halted in recent years as they have shown limited therapeutic benefit. This is in part due to the considerable heterogeneity in the disease course and the fact that treatments are likely delivered too late due to the preclinical lag in symptoms. In the context of drugs to target the epigenome, the matter is further complicated by the issues of causality, targeting the correct cell type and the dynamicity of epigenetic marks. Nonetheless, epigenetic drugs do represent considerable clinical promise for AD and movements are being made towards developing these treatments. The Biopharmaceutical company

Oryzon currently has a Phase IIa clinical trial 'Epigenetic Therapy in Alzheimer's disease (ETHERAL)', which is evaluating the safety and tolerability of Vafidemstat, a dual lysine-specific histone demethylase, in patients with mild-to-moderate AD [65]. At present, no compounds are being pursued in AD clinical trials that specifically target the DNA methylation machinery. However, we hope that as a result of the robust findings emerging from DNA methylomic studies of AD, combined with advances in pharmaceutical technologies, this will simply be a matter of time.

Data availability statement

Data sharing not applicable to this article as no data sets were generated or analysed during the current study.

References

- 1 Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006; **31**: 89–97
- 2 Alzheimer's Association. Alzheimer's Disease Facts and Figures. *Alzheimer's and Dementia* 2017; **13**: 325–73
- 3 Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*. 2007; **29**: 125–32
- 4 Watt F, Molloy PL. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* 1988; **2**: 1136–43
- 5 Prince M, Anders W, Maëllenn G, Gemma-Claire A, Yu-Tzu W, Matthew P; World Alzheimer Report. The Global Impact of Dementia, an analysis of prevalence, incidence, cost and trends. *Alzheimer's Disease Int* 2015; **2015**: 1–87
- 6 Miller JL, Grant PA. The role of DNA methylation and histone modifications in transcriptional regulation in humans. *Sub-Cellular Biochem* 2013; **61**: 289–317
- 7 Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological Alterations in Alzheimer Disease. *Cold Spring Harb Perspect Med* 2011; **1**: a006189
- 8 Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, et al. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res* 2013; **23**: 555–67
- 9 Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 1991; **82**: 239–59
- 10 Scheff SW, DeKosky ST, Price DA. Quantitative assessment of cortical synaptic density in Alzheimer's disease. *Neurobiol Aging* 1990; **11**: 29–37
- 11 Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res* 2013; **23**: 1256–69
- 12 Gomez-Isla T, Price JL, McKeel DW Jr, Morris JC, Growdon JH, Hyman BT. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J Neurosci: The Official Journal of the Society for Neuroscience* 1996; **16**: 4491–500
- 13 Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 2010; **38**: e181
- 14 Meda L, Baron P, Scarlato G. Glial activation in Alzheimer's disease: the role of Abeta and its associated proteins. *Neurobiol Aging* 2001; **22**: 885–93
- 15 Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009; **324**: 929–30
- 16 Khare T, Pai S, Koncivicius K, Pal M, Kriukiene E, Liutkeviciute Z, et al. 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nat Struct Mol Biol* 2012; **19**: 1037–43
- 17 Aisen PS, Cummings J, Jack CR, Morris JC, Sperling R, Frölich L, et al. On the path to 2025: understanding the Alzheimer's disease continuum. *Alzheimer's Res Ther* 2017; **9**: 60
- 18 Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 2006; **63**: 168–74
- 19 Campion D, Dumanchin C, Hannequin D, Dubois B, Belliard S, Puel M, et al. Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *Am J Hum Genet* 1999; **65**: 664–70
- 20 Sadakierska-Chudy A, Filip M. A Comprehensive View of the Epigenetic Landscape. Part II: Histone Post-translational Modification, Nucleosome Level, and Chromatin Regulation by ncRNAs. *Neurotox Res* 2015; **27**: 172–97
- 21 Bekris LM, Yu C-E, Bird TD, Tsuang DW. Review Article: Genetics of Alzheimer Disease. *J Geriatr Psychiatry Neurol* 2010; **23**: 213–27
- 22 Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* 2019; **51**: 404–13
- 23 Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and

- implicates A β , tau, immunity and lipid processing. *Nat Genet* 2019; **51**: 414–30
- 24 Kim YZ. Altered Histone Modifications in Gliomas. *Brain Tumor Res Treat*. 2014; **2**: 7–21
 - 25 Michaelson DM. APOE ϵ 4: The most prevalent yet understudied risk factor for Alzheimer's disease. *Alzheimer's Dement*. 2014; **10**: 861–8
 - 26 Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 2018; **47**: D155–62
 - 27 Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* 2004; **116**: 281–97
 - 28 Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain: a Journal of Neurology*. 2015; **138**: 3673–84
 - 29 Ridge PG, Mukherjee S, Crane PK, Kauwe JS. Alzheimer's Disease Genetics C. Alzheimer's disease: analyzing the missing heritability. *PLoS One* 2013; **8**: e79771
 - 30 Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 Variants in Alzheimer's Disease. *N Engl J Med* 2013; **368**: 117–27
 - 31 Park JS, Lee J, Jung ES, Kim M-H, Kim IB, Son H, et al. Brain somatic mutations observed in Alzheimer's disease associated with aging and dysregulation of tau phosphorylation. *Nat Commun* 2019; **10**: 3090
 - 32 Henikoff S, Matzke MA. Exploring and explaining epigenetic effects. *Editorial* 1997; **13**: 293–5
 - 33 Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell* 2009; **136**: 215–33
 - 34 Roubroeks JAY, Smith RG, van den Hove DLA, Lunnon K. Epigenetics and DNA methylomic profiling in Alzheimer's disease and other neurodegenerative diseases. *J Neurochem* 2017; **143**: 158–70
 - 35 Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci* 2014; **17**: 1164–70
 - 36 De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci* 2014; **17**: 1156–63
 - 37 Lord J, Cruchaga C. The epigenetic landscape of Alzheimer's disease. *Nat Neurosci* 2014; **17**: 1138–40
 - 38 Smith RG, Hannon E, De Jager PL, Chibnik L, Lott SJ, Condliffe D, et al. Elevated DNA methylation across a 48-kb region spanning the HOXA gene cluster is associated with Alzheimer's disease neuropathology. *Alzheimer's Dement* 2018; **14**: 1580–8
 - 39 Watson CT, Roussos P, Garg P, Ho DJ, Azam N, Katsel PL, et al. Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. *Genome Med* 2016; **8**: 5
 - 40 Mano T, Nagata K, Nonaka T, Tarutani A, Imamura T, Hashimoto T, et al. Neuron-specific methylome analysis reveals epigenetic regulation and tau-related dysfunction of BRCA1 in Alzheimer's disease. *Proc Natl Acad Sci USA* 2017; **114**: E9645–e54
 - 41 Gasparoni G, Bultmann S, Lutsik P, Kraus TFJ, Sordon S, Vlcek J, et al. DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex. *Epigenetics & Chromatin*. 2018; **11**: 41
 - 42 Smith AR, Smith RG, Pishva E, Hannon E, Roubroeks JAY, Burrage J, et al. Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation in Alzheimer's disease. *Clinical epigenetics*. 2019; **11**: 52
 - 43 Lardenoije R, Roubroeks JAY, Pishva E, Leber M, Wagner H, Iatrou A, et al. Alzheimer's disease-associated (hydroxy)methylomic changes in the brain and blood. *Clinical epigenetics*. 2019; **11**: 164
 - 44 Zhao J, Zhu Y, Yang J, Li L, Wu H, De Jager PL, et al. A genome-wide profiling of brain DNA hydroxymethylation in Alzheimer's disease. *Alzheimer's Dement* 2017; **13**: 674–88
 - 45 Ellison EM, Bradley-Whitman MA, Lovell MA. Single-Base Resolution Mapping of 5-Hydroxymethylcytosine Modifications in Hippocampus of Alzheimer's Disease Subjects. *Journal of molecular neuroscience: MN* 2017; **63**: 185–97
 - 46 Smith RG, Pishva E, Shireby G, Smith AR, Roubroeks JAY, Hannon E, et al. Meta-analysis of epigenome-wide association studies in Alzheimer's disease highlights novel differentially methylated loci across cortex. *BioRxiv* 2020; <https://doi.org/10.1101/2020.02.28.957894>
 - 47 Smith AR, Smith RG, Burrage J, Troakes C, Al-Sarraj S, Kalaria RN, et al. A cross-brain regions study of ANK1 DNA methylation in different neurodegenerative diseases. *Neurobiol Aging* 2019; **74**: 70–6
 - 48 Andrade-Moraes CH, Oliveira-Pinto AV, Castro-Fonseca E, da Silva CG, Guimaraes DM, Szczupak D, et al. Cell number changes in Alzheimer's disease relate to dementia, not to plaques and tangles. *Brain: a journal of neurology*. 2013; **136**: 3738–52
 - 49 Hernandez HG, Sandoval-Hernandez AG, Garrido-Gil P, Labandeira-Garcia JL, Zelaya MV, Bayon GF, et al. Alzheimer's disease DNA methylome of pyramidal layers in frontal cortex: laser-assisted microdissection study. *Epigenomics*. 2018; **10**: 1365–82
 - 50 Karemaker ID, Vermeulen M. Single-Cell DNA Methylation Profiling: Technologies and Biological Applications. *Trends Biotechnol* 2018; **36**: 952–65
 - 51 Olova N, Krueger F, Andrews S, Oxley D, Berrens RV, Branco MR, et al. Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data. *Genome Biol* 2018; **19**: 33

- 52 English AC, Richards S, Han Y, Wang M, Vee V, Qu J, et al. Mind the Gap: Upgrading Genomes with Pacific Biosciences RS Long-Read Sequencing Technology. *PLoS One* 2012; 7: e47768
- 53 Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 2010; 7: 461–5
- 54 Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, et al. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc Natl Acad Sci USA* 2013; 110: 18904–9
- 55 Song CX, He C. Potential functional roles of DNA demethylation intermediates. *Trends Biochem Sci* 2013; 38: 480–4
- 56 Wu H, Wu X, Zhang Y. Base-resolution profiling of active DNA demethylation using MAB-seq and caMAB-seq. *Nat Protoc* 2016; 11: 1081
- 57 Fetaahu IS, Ma D, Rabidou K, Argueta C, Smith M, Liu H, et al. Epigenetic signatures of methylated DNA cytosine in Alzheimer's disease. *Science. Advances* 2019; 5: eaaw2880
- 58 Wu X, Inoue A, Suzuki T, Zhang Y. Simultaneous mapping of active DNA demethylation and sister chromatid exchange in single cells. *Genes Dev* 2017; 31: 511–23
- 59 Relton CL, Davey SG. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol* 2012; 41: 161–76
- 60 Relton CL, Davey Smith G. Mendelian randomization: applications and limitations in epigenetic studies. *Epigenomics* 2015; 7: 1239–43
- 61 Imm J, Kerrigan TL, Jeffries A, Lunnon K. Using induced pluripotent stem cells to explore genetic and epigenetic variation associated with Alzheimer's disease. *Epigenomics* 2017; 9(11): 1455–68
- 62 Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, et al. Editing DNA Methylation in the Mammalian Genome. *Cell* 2016; 167(1): 233–47.e17
- 63 Qian X, Song H, Ming G-l. Brain organoids: advances, applications and challenges. *Development* 2019; 146: dev166074
- 64 Saito T, Mihira N, Matsuba Y, Sasaguri H, Hashimoto S, Narasimhan S, et al. Humanization of the entire murine Mapt gene provides a murine model of pathological human tau propagation. *The Journal of biological chemistry* 2019; 294: 12754–65
- 65 Maes T, Mascaro C, Ortega A, Lunardi S, Ciceri F, Somervaille TC, et al. KDM1 histone lysine demethylases as targets for treatments of oncological and neurodegenerative disease. *Epigenomics* 2015; 7: 609–26

Received 4 October 2019

Accepted after revision 12 July 2020

Published online Article Accepted on 3 August 2020